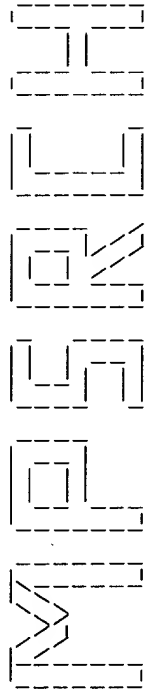


\*\*\*\*\*



(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

on: Wed Jul 22 15:27:41 1998; MasPar time 12.38 Seconds  
767.369 Million cell updates/sec  
Tabular output not generated.

Title: >US-08-915-659A-7  
Description: (1-260) from US08915659A.pep  
Perfect Score: 1943  
Sequence: 1 MGRPRPRAAKTWMLLLGG.....VYTNICRYLDWIKKIIGSKG 260

Scoring table: PAM 150  
Gap 11

Searched: 120441 seqs, 36531193 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: p156  
1:p1r1 2:p1r2 3:p1r3 4:p1r4 5:n13d

Statistics: Mean 42.128; Variance 70.703; scale 0.596

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Query	Score	Match	Length	DB	ID	Description	Pred. No.
1	1488	76.6	260	2	I56559	neuropsin - mouse	0.00e+00
2	833	42.9	253	2	A53968	serine proteinase SCC	3.84e-170
3	808	41.6	247	1	TRDG	trypsin (EC 3.4.21.4)	5.44e-164
4	804	41.4	238	2	S31779	trypsin (EC 3.4.21.4)	5.23e-163
5	795	40.9	246	1	TRDGC	trypsin (EC 3.4.21.4)	8.52e-161
6	785	40.4	229	1	TRBOPR	trypsin (EC 3.4.21.4)	2.43e-158
7	783	40.3	247	2	A27547	trypsin (EC 3.4.21.4)	7.53e-158
8	781	40.2	246	1	TRRT2	trypsin (EC 3.4.21.4)	2.33e-157
9	779	40.1	225	5	ITGSZ	trypsin (EC 3.4.21.4)	7.22e-157
10	768	39.5	223	5	2TIDE	trypsin (EC 3.4.21.4)	3.60e-154
11	768	39.5	246	2	B25528	trypsin (EC 3.4.21.4)	3.60e-154
12	767	39.5	246	1	TRRT1	trypsin (EC 3.4.21.4)	6.33e-154
13	766	39.4	223	5	ITPS	trypsin complexed wit	1.11e-153
14	766	39.4	223	5	ITYN	Beta trypsin complex	1.11e-153
15	766	39.4	223	5	IURTA	trypsin (EC 3.4.21.4)	1.11e-153
16	766	39.4	223	5	IURSA	trypsin (EC 3.4.21.4)	1.11e-153
17	766	39.4	223	5	ITGB	trypsin (EC 3.4.21.4)	1.11e-153
18	766	39.4	223	5	ITGT	trypsin (EC 3.4.21.4)	1.11e-153
19	766	39.4	223	5	4PTPIZ	trypsin (EC 3.4.21.4)	1.11e-153
20	766	39.4	223	5	2PTN	trypsin (EC 3.4.21.4)	1.11e-153
21	766	39.4	223	5	ITPP	trypsin (EC 3.4.21.4)	1.11e-153
22	766	39.4	223	5	ITBO	trypsin (EC 3.4.21.4)	1.11e-153
23	766	39.4	223	5	IBTY	Beta-trypsin (EC 3.4.21.4)	1.11e-153

ALIGNMENTS

```
RESULT 1
ENTRY      I56559      #type complete
TITLE      neuropsin - mouse
ORGANISM   Mus musculus #common_name house mouse
DATE       26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change
ACCESSION  I56559
REFERENCE  Ched Z.L.; Yoshida, S.; Kato, K.; Momota, Y.; Suzuki, J.;
            Tanaka, T.; Ito, J.; Nishino, H.; Aimoto, S.; Kiyama, H.;
            J. Neurosci. (1995) 15:5088-5097
            Expression and activity-dependent changes of a novel
            limbic-serine protease gene in the hippocampus.
#cross-references MUID:95348817
#accession I56559
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 1-260 #label RES
#cross-references GB:D30785; NID:gl020090; PID:gl020091
CLASSIFICATION #superfamily trypsin; trypsin homology
FEATURE
33-252      #domain trypsin homology #label TRY
SUMMARY     #length 260 #molecular-weight 28523 #checksum 6448
```

```
Query Match 76.6%; Score 1488; DB 2; Length 260;
Best Local Similarity 72.2%; Pred. No. 0.00e+00;
Matches 187; Conservative 38; Mismatches 34; Indels 0; Gaps 0;

Db 1 MGRPPCAIQPWILLFMGAWAGLTRAQGSKILEGRECIPIHSPQWQAALFOGERLICGG 60
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 1 MGRPRPRAAKTWMLLLGGAWAGHSAQEDKVLGHECQPHSQWQAALFOGQQLCGG 60
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

Db 61 VLYGDRWLTAACHCKKQKYSVRLGDSLSRQDPEBIEIQAQSIQHPCTYNNPDEHSD 120
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 61 VLYGGNWLTAACHCKPKYKTVRLGDSLSLQNKDGPEDIEPVVQSIPIHPCTYNNSSVEDHNHD 120
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

Db 121 IMILRLQNSANLGDVKPKVOLANLCPKVGOKCIISGWGTVTSPQENFPNTLCAEYKIYS 180
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 121 LMLQLRQDLSKSLGKPKYKSLADHCTQPGKCTVSGWGTVTSPRENFPTLCAEYKIFP 180
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

Db 181 QNKCEAYPGKITEGVMVCAAGSSNGADTCQDSSGGLVDCGMLQGITSWGSDPGCKPEKPG 240
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 181 QKCEDAYPGQITDGMVCAAGSSKGDTCQDSSGGLVDCGALQGITSWGSDPGCRSDKPG 240
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

Db 241 VYTKICRYTWIKTKWMDNR 259
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 241 VYTNICRYLDWIKKIIGSK 259
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
```

```

2
RESULT 2
ENTRY A53968 #type complete
TITLE serine proteinase SCCE precursor - human
ALTERNATE_NAMES stratum corneum chymotryptic enzyme
ORGANISM #formal_name Homo sapiens #common_name man
DATE 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change
08-Sep-1997

ACCESSIONS A53968
REFERENCE A53968
#authors Hansson, L.; Stroemqvist, M.; Baeckman, A.; Wallbrandt, P.;
#journal Carlstein, A.; Egelrud, T.;
#title J. Biol. Chem. (1994) 269:19420-19426
#disulfide_bonds #status predicted
#active_site His, Asp, Ser #status predicted
#binding_site calcium (Glu, Asn, Val, Glu) #status
#predicted
#length 247 #molecular-weight 27525 #checksum 644

SUMMARY
Query Match 42.9%; Score 833; DB 2; Length 253;
Best Local Similarity 47.3%; Pred. No. 3.84e-170;
Matches 116; Conservative 44; Mismatches 77; Indels 8; Gaps 6;

Db 11 ILLLSLALETAGEE-AQGDKIIDAPCARGHPQVALLSQNLHCGGLVNERWVLTAA 69
QY 14 FLLL-LGGAWAGHRAQEDKVLGGHECQPHSQPQWQALFQGGQLCGGLVGNWVLTAA 72
Db 70 HCKMNEYTVHLGSDTLGDR--RAQRIKASKSRFRPGY-ST--QTHVNDLMLVKNLSQARL 124
QY 73 HCKPKYTVRLGSHLQNKQDGEPIPVQSIHPHCYNSSDVEDHNDLMLQLRDOASL 132
Db 125 SSMVKVRLPSRCPPGTTCTVSWGTTTSPDVTTPSDLMCDVVKLISPDQTKVYKDLL 184
QY 133 GSKVKPISLADHCTQPGQKCTVSGWGTVTSPRENFPDTLNCAEVKIFPQKCEDAYPGQI 192
Db 185 ENSMLCAGIPDSKKNAGDSGGPLVCRGTGLGVSMGTFCGPNPDGVTQVCKETKW 244
QY 193 TDGMVACGSSKGA-DTCOGDSGGPLVCDGALQGITSWGSDPCGRSDRPGVYTNICRYLDW 251
Db 245 INDTM 249
QY 252 IKKI 256

RESULT 3
ENTRY TRDG #type complete
TITLE trypsin (EC 3.4.21.4) precursor, anionic - dog
ALTERNATE_NAMES cationic trypsinogen
ORGANISM #formal_name Canis lupus familiaris #common_name dog
DATE 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change
18-Jul-1997

ACCESSIONS A26273
REFERENCE A26273
#authors Pinsky, S.D.; LaForge, K.S.; Scheele, G.
#journal Mol. Cell. Biol. (1985) 5:2669-2676
#title Differential regulation of trypsinogen mRNA translation:
#full-length mRNA sequences encoding two oppositely charged
#trypsinogen isoenzymes in the dog pancreas.
#cross-references MUID:86284628
#accession A26273
#molecule_type mRNA

```

```

##residues 1-247 ##label PIN
CLASSIFICATION #superfamily trypsin; trypsin homology
KEYWORDS hydrolase; pancreas; protein digestion; serine proteinase;
zymogen
FEATURE
1-15 #domain signal sequence #status predicted #label SIG\
16-23 #domain activation peptide #status predicted #label AP\
24-247 #product trypsin, anionic #status predicted #label ENZ\
24-239 #domain trypsin homology #label TRY\
30-160,48-64,
132-233,139-206,
171-185 #disulfide_bonds #status predicted\
63,107,200 #active_site His, Asp, Ser #status predicted\
75,77,80,85 #binding_site calcium (Glu, Asn, Val, Glu) #status
#predicted
SUMMARY #length 247 #molecular-weight 26423 #checksum 8431

Query Match 41.6%; Score 808; DB 1; Length 247;
Best Local Similarity 44.2%; Pred. No. 5.44e-164;
Matches 110; Conservative 57; Mismatches 75; Indels 7; Gaps 7;

Db 5 LILAFLGAA-VATPTDDDDKIVGGYTCENSVPVQVSLNAGYHF-CGGSLISDQWVVSAA 62
QY 13 MFLLLLGGAWAGHRAQEDKVLGGHECQPHSQPQWQALFQGGQLCGGLVGNWVLTAA 72
Db 63 HCYSRIQVRLGEYNIDVLEGNQFINSAKVIIRHPNYSW-ILD-N-DIMLIKLSPPAVL 119
QY 73 HCKPKYTVRLGSHLQNKQDGEPIPVQSIHPHCYNSSDVEDHNDLMLQLRDOASL 132
Db 120 NARVAISLPRAACAPGTQCLISGWNTLSSGTYNPELLOCLDAPILTOACASYPGOI 179
QY 133 GSKVKPISLADHCTQPGQKCTVSGWGTVTSPRENFPDTLNCAEVKIFPQKCEDAYPGQI 192
Db 180 TENMICAGFLEGKDCSGDSGGPVVNCNGELQIVSGYIG-CAQKNKPGVYTVKCNFVDW 238
QY 193 TDGMVACGSSKGA-DTCOGDSGGPLVCDGALQGITSWGSDPCGRSDRPGVYTNICRYLDW 251
Db 239 IQSTIAANS 247
QY 252 IKKI 256

RESULT 4
ENTRY S31779 #type fragment
TITLE trypsin (EC 3.4.21.4) III precursor - Atlantic salmon
ORGANISM #formal_name Salmo salar #common_name Atlantic salmon
DATE 22-Nov-1993 #sequence_revision 03-Aug-1995 #text_change
07-Nov-1997
ACCESSIONS S66657; S31779
REFERENCE S66657
#authors Male, R.; Lorens, J.B.; Smalas, A.O.; Torrisen, K.R.
#journal Eur. J. Biochem. (1995) 232:677-685
#title Molecular cloning and characterization of anionic and
#cationic variants of trypsin from Atlantic salmon.
#accession S66657
#molecule_type mRNA
##residues 1-238 ##label MAL
##cross-references EMBL:X70074; NID:964387; PID:964388
CLASSIFICATION #superfamily trypsin; trypsin homology
KEYWORDS hydrolase; serine proteinase
FEATURE
1-7 #domain signal sequence (fragment) #status predicted
#label SIG\
8-15 #domain activation peptide #status predicted #label AP\
16-238 #product trypsin III #status predicted #label MAR\
16-231 #domain trypsin homology #label TRY\
22-152,40-56,
124-225,131-198,
163-177,188-212
55,99,192 #disulfide_bonds #status predicted\
#active_site His, Asp, Ser #status predicted
SUMMARY #length 238 #checksum 6737

```

QY	190	GOITDMWCASSKGA-DTCQDGGPLVCDGALQGITSWGSDPCGRSKDPGYVYINCRY	248
Db	236	VSWIOQTIAAN	246
QY	249	LDWIKKLIISK	259
RESULT	6		
ENTRY		TRBOTR	#type complete
TITLE		trypsin (EC 3.4.21.4) precursor - bovine	
CONTAINS		trypsinogen	
ORGANISM		#formal_name Bos primigenius taurus	#common_name cattle
DATE		24-Apr-1984	#sequence_revision 28-Feb-1986
		18-Jul-1997	#text_change
ACCESSIONS		A90164; A00946; S08774	
REFERENCE		A90164	
#authors		Mikes, O.; Holeysovsky, V.; Tomasek, V.; Sorm, F.	
#journal		Biochem. Biophys. Res. Commun. (1966) 24:346-352	
#title		Covalent structure of bovine trypsinogen. The position of the remaining amides.	
#cross-references		MUID:67168848	
#accession		A90164	
#molecule_type		protein	
#residues		1-57,'Q','59-67','Q','69-150','N','152-176','N','178-229	
		##label MKK	
REFERENCE		A93755	
#authors		Hartley, B.S.	
#journal		Philos. Trans. R. Soc. Lond. (1970) B257:77-87	
#contents		annotation; revisions	
REFERENCE		A00950	
#authors		Titani, K.; Ericsson, L.H.; Neurath, H.; Walsh, K.A.	
#journal		Biochemistry (1975) 14:1358-1366	
#title		Amino acid sequence of dogfish trypsin.	
#cross-references		MUID:7514645	
#contents		annotation; revisions	
#note		the sequence agrees with that shown	
REFERENCE		A92954	
#authors		Bode, W.; Schwager, P.	
#journal		J. Mol. Biol. (1975) 98:693-717	
#title		The refined crystal structure of bovine beta-trypsin at 1.8 angstrom resolution.	
#cross-references		MUID:76072097	
#contents		annotation; X-ray crystallography; binding sites for calcium, substrate, and inhibitors	
COMMENT		Trypsinogen is synthesized in the acinar cells of the pancreas.	
COMMENT		Autocatalytic cleavage after Lys-6 leads to beta-trypsin by releasing a terminal hexapeptide. Subsequent cleavage after Lys-131 leads to alpha-trypsin. Further cleavage after Lys-176 yields pseudotrypsin. A cleavage may also occur after Arg-105. #superfamily trypsin; trypsin homology	
CLASSIFICATION		hydrolase; pancreas; protein digestion; serine proteinase; zymogen	
KEYWORDS			
FEATURE			
1-229		#product trypsinogen	#status experimental
1-6		#domain activation peptide	#status experimental
		APV	
7-222		#domain trypsin homology	#label TRV
7-131,132-229		#product alpha-trypsin	#status experimental
6-7		#cleavage_site Lys-Ile (enteropeptidase)	#status experimental
13-143,31-47, 115-216,122-189, 154-168,179-203, 46.90.183 58.60.63.68		#disulfide_bonds	#status experimental
		#active_site His, Asp, Ser	#status experimental
		#binding_site calcium (Glu, Asn, Val, Glu)	#status experimental
131-132		#cleavage_site Lys-Ser (autolytic)	#status experimental
SUMMARY		#length 229	#molecular-weight 23993
			#checksum 2248
Query Match		40.4%	Score 785; DB 1; Length 229;
Best Local Similarity		46.6%	Pred. No. 2.43e-158;
Matches 108; Conservative		50; Mismatches 58;	Indels 6; Gaps 6;

```

Db 3 DDDKIVGGYTCGANTVPVQVSLNSGYHF-CGSLINSQWVYVAAHCKYSGIQVRLGEDNI 61
Qy 29 QEDRVLGHECHQSPQWQAALFOGQQLLGGVVLGGVNWVLAHCKKPKYTVRLGDHSL 88
Db 62 NVVSGNQFISAKSIYHPSYNSNTL-N-N-DIMLIKLSAASLSNRVASISLPTSCASA 118
Qy 89 QNKDGPQEIIPVQSIHPCYNSSDVEDHNDMLLQDQASLGSKVKPISLADHCTQP 148
Db 119 GTQCLISGWNKTSKSGTSYDVLKCLAPILSDSCSKSAYPGQITSNMFCAGYEGGKDS 178
Qy 149 GQKCTVSGWGTVPSPRENFPTLCAEVKIPQKCEDAYPGQITDGMVCAAGSKGA-DT 207
Db 179 CQSDSGPVVCSGKLGQIVSGSG-CAQKNKPGYVTVKCNVSVIKOTIASN 229
Qy 208 CQSDSGPLVCDGALQGITSGSDPCGRSDKPGVYTNICRYLDWIKKIIGSK 259

RESULT 7
ENTRY trypsin (EC 3.4.21.4) precursor, cationic - rat
TITLE trypsin (EC 3.4.21.4) precursor, cationic - rat
ORGANISM Rattus norvegicus #common_name Norway rat
#formal_name Rattus norvegicus #sequence_revision 30-Jun-1988 #text_change
20-Mar-1998
ACCESSIONS A27547
REFERENCE A27547
#authors Craik, T.S.; Alhadeff, M.; Craik, C.S.; Largman, C.
#journal Biochemistry (1987) 26:3081-3086
#title Isolation and characterization of a cDNA encoding rat
#cross-references MUID:8721609
#accession A27547
#molecule_type mRNA
#residues 1-247 #label FILE
#cross-references GB:M16624; NID:Q206498; PID:Q206499
CLASSIFICATION #superfamily trypsin; trypsin homology
KEYWORDS calcium binding; hydrolase; protein digestion; serine
proteinase
FEATURE
25-240 #domain trypsin homology #label TRY\
31-161,49-65, #disulfide_bonds #status predicted\
133-234,140-207, #active_site His, Asp, Ser #status predicted\
172-186 #binding_site calcium (Glu, Asn, Val, Glu) #status
64,108,201 predicted
76,78,81,86 #length 247 #molecular-weight 26269 #checksum 9564
SUMMARY
Query Match 40.3%; Score 783; DB 2; Length 247;
Best Local Similarity 43.0%; Pred. No. 7.53e-158;
Matches 108; Conservative 55; Mismatches 82; Indels 6; Gaps 6;
Db 2 KALIFLAFGLAAVALPLDDDDKIVGGYTCQKNSLPYQVSLNAGYHF-CGSLINSQWVY 60
Qy 10 KTWFWLLGGWAGHRAQEDKVLGHECHQSPQWQAALFOGQQLLGGVVLGGVNWV 69
Db 61 SAACHYKSIQVRLGEHNIYVGEQFIDAAKILIRHPSYNTAF-D-N-DIMLIKLSNP 117
Qy 70 TAAHCKPKYTVRLGDHSLQNKDGPQEIIPVQSIHPCYNSSDVEDHNDMLLQDQ 129
Db 118 ATLSNRVSTVSLPRSCSGSKLVSGWNTLSGTYNPSLQCLDAPVLSDSCKSSYP 177
Qy 130 ASLGSVKPISLADHCTQPGQKCTVSGWGTVPSPRENFPTLCAEVKIPQKCEDAYP 189
Db 178 GKITSNMFCFLGLEGKDSQCGSDSGPVVCSGKLGQIVSGSG-CAQKNKPGYVTVKCN 236
Qy 190 QGITDGMVCAAGSKGA-DTCQSDGGLVCDGALQGITSGSDPCGRSDKPGVYTNICRY 248
Db 237 VNWIQOTVAAN 247
Qy 249 LDWIKKIIGSK 259

RESULT 8
ENTRY trypsin (EC 3.4.21.4) precursor (with pancreatic secretory

```

```

ENTRY trypsin (EC 3.4.21.4) II precursor - rat
TITLE trypsinogen II
ALTERNATE_NAMES #formal_name Rattus norvegicus #common_name Norway rat
ORGANISM Rattus norvegicus #sequence_revision 30-Sep-1987 #text_change
05-Apr-1983
DATE 18-Jul-1997
ACCESSIONS A22657; A00949
REFERENCE A22657
#authors Craik, C.S.; Choo, O.L.; Swift, G.H.; Quinto, C.; MacDonald,
R.J.; Rutter, W.J.
#journal J. Biol. Chem. (1984) 259:14255-14264
#title Structure of two related rat pancreatic trypsin genes.
#cross-references MUID:85054880
#accession A22657
#molecule_type DNA
#residues 1-246 #label CRA
REFERENCE A00948
#authors MacDonald, R.J.; Stary, S.J.; Swift, G.H.
#journal J. Biol. Chem. (1982) 257:9724-9732
#title Two similar but nonallelic rat pancreatic trypsinogens.
#cross-references MUID:82265624
#accession A00949
#molecule_type mRNA
#residues 9-246 #label MAC
COMMENT The trypsin II mRNA is present in much lower quantities than the
trypsin I mRNA.
GENETICS 14/1; 67/2
#introns
CLASSIFICATION #superfamily trypsin; trypsin homology
KEYWORDS hydrolase; pancreas; protein digestion; serine proteinase;
zymogen
FEATURE
1-15 #domain signal sequence #status predicted #label SIG\
16-23 #domain activation peptide #status predicted #label AP\
24-246 #product trypsin II #status predicted #label ENZ\
24-239 #domain trypsin homology #label TRY\
30-160,48-64, #disulfide_bonds #status predicted\
132-233,139-206, #active_site His, Asp, Ser #status predicted\
171-185 #binding_site calcium (Glu, Asn, Val, Glu) #status
63-107,200 predicted
75,77,80,85 #length 246 #molecular-weight 26243 #checksum 3816
SUMMARY
Query Match 40.2%; Score 781; DB 1; Length 246;
Best Local Similarity 42.6%; Pred. No. 2.33e-157;
Matches 107; Conservative 56; Mismatches 81; Indels 7; Gaps 7;
Db 2 RALLFLALGAA-VAPFVDDDDKIVGGYTCQENSVPYQVSLNSGYHF-CGSLINSQWVY 59
Qy 10 KTWFWLLGGWAGHRAQEDKVLGHECHQSPQWQAALFOGQQLLGGVVLGGVNWV 69
Db 60 SAACHYKSIQVRLGEHNIYVGEQFIDAAKILIRHPSYNTAF-D-N-DIMLIKLSNP 116
Qy 70 TAAHCKPKYTVRLGDHSLQNKDGPQEIIPVQSIHPCYNSSDVEDHNDMLLQDQ 129
Db 117 VKLNARVAVLPSKCAPAGTQCLISGWNLSGTYNPSLQCLDAPVLSDSCKSSYP 176
Qy 130 ASLGSVKPISLADHCTQPGQKCTVSGWGTVPSPRENFPTLCAEVKIPQKCEDAYP 189
Db 177 GKITDNMVCFLGLEGKDSQCGSDSGPVVCSGKLGQIVSGSG-CALPDNPGYVTVKCN 235
Qy 190 QGITDGMVCAAGSKGA-DTCQSDGGLVCDGALQGITSGSDPCGRSDKPGVYTNICRY 248
Db 236 VDWIQOTVAAN 246
Qy 249 LDWIKKIIGSK 259

RESULT 9
ENTRY trypsin (EC 3.4.21.4) precursor (with pancreatic secretory

```

```

trypsin inhibitor), chain Z - bovine
#formal_name Bos primigenius taurus #common_name cattle
#note pancreas
REFERENCE
#authors
Bolognesi, M.; Gatti, G.; Menegatti, E.; Guarneri, M.;
Marquart, M.; Papamokos, E.; Huber, R.
#submission submitted to the Brookhaven Protein Data Bank, September 1982
#cross-references PDB:1UGS
REFERENCE
#authors
TN005560
Bolognesi, M.; Gatti, G.; Menegatti, E.; Guarneri, M.;
Marquart, M.; Papamokos, E.; Huber, R.
#journal
J. Mol. Biol. (1982) 162:839
#title
Three-dimensional structure of the complex between pancreatic
secretory inhibitor (kazal type) and trypsinogen at 1.8
angstroms resolution. structure solution, crystallographic
refinement and preliminary structural interpretation.
TN005561
Marquart, M.; Walter, J.; Deisenhofer, J.; Bode, W.; Huber,
R.
Acta Crystallogr. (1983) B39:480
The geometry of the reactive site and of the peptide groups
in trypsin, trypsinogen and its complexes with inhibitors.
TN005562
Dayhoff, M.O.
in Atlas of Protein Sequence and Structure (Data Section),
Dayhoff, M.O., ed., vol. 5, pp.105, National Biomedical
Research Foundation, Silver Spring, Md., 1972
TN005563
Dayhoff, M.O.
in Atlas of Protein Sequence and Structure, Supplement 1,
Dayhoff, M.O., ed., vol. 5, pp.88, National Biomedical
Research Foundation, Silver Spring, Md., 1973
COMMENT Resolution: 1.8 angstroms
COMMENT R-value: 0.186
COMMENT Determination: X-ray diffraction
KEYWORDS hydrolase; serine proteinase
FEATURE
146-158
#region helix (right hand alpha) (sngl alpha turn, rest
irreg.)\
#region helix (right hand alpha) (contiguous with h3)\
#region helix (right hand alpha) (contiguous with h2)\
#disulfide_bonds\
#disulfide_bonds\
#disulfide_bonds\
#disulfide_bonds\
#disulfide_bonds\
#disulfide_bonds\
#length 225 #molecular-weight 23548 #checksum 5415
Query Match 40.1%; Score 779; DB 5; Length 225;
Best Local Similarity 47.0%; Pred. No. 7.22e-157;
Matches 108; Conservative 48; Mismatches 68; Indels 6; Gaps 6;
Db 1 DKIVGYTCGANTVPQVLSNGYHF-CGGSLINSQWVVAHCKYKGIQVRLGDNINV 59
Qy 31 DKVLGGHCQPHSQFQWQAALFQGLLQGLVGNWVLTAAHCKPKYTVRLGDHSLQN 90
Db 60 VEGNEQFISAKSIIVHPSYNSNTL-N-DIMLIKLSAASLNSRVASISLPTSCASAGT 116
Qy 91 KGPPEQIIPVQSIHPHCYNSSDVEDHNDMLQLRDQASLGSVKVPISLADHCTOPGQ 150
Db 117 QCLISGWGNTKSSGTSPDVLKCLKAPILSDSSCKSAVPQGITSNMFCAGYLEGGKDSQC 176
Qy 151 KCTVSGWGTVPSPRENFDTLNCAEVKIFPKQKCEDAYPGQITDGMVCAGSKGA-DTCQ 209
Db 177 GDSGGPVVCSKLGIVSWGSG-CAQKNKPGVTKVCNYSWIKOTIASN 225
Qy 210 GDSGGLVCDGALQGITSWGSDPCGRSKPGVYTNICRYLDWIKKIIGSK 259
RESULT 10
ENTRY 2TLDE #type complete
TITLE trypsin (EC 3.4.21.4) (with mutant streptomyces subtilisin

```

```

inhibitor), chain E - bovine
#formal_name Bos primigenius taurus #common_name cattle
#note pancreas
REFERENCE
#authors
Mitsui, Y.; Takeuchi, Y.; Nonaka, T.; Nakamura, K.T.
submitted to the Brookhaven Protein Data Bank, September 1991
#cross-references PDB:2TLD
REFERENCE
#authors
TN005550
Takeuchi, Y.; Nonaka, T.; Nakamura, K.T.; Kojima, S.; Miura,
K.I.; Mitsui, Y.
Proc. Natl. Acad. Sci. U.S.A. (1992) 89:4407
Crystal structure of an engineered subtilisin inhibitor
complexed with bovine trypsin.
TN005551
Takeuchi, Y.; Satow, Y.; Nakamura, K.T.; Mitsui, Y.
J. Mol. Biol. (1991) 221:309
Refined crystal structure of the complex of subtilisin BPN'
and streptomyces subtilisin inhibitor at 1.8 angstroms
resolution.
COMMENT Resolution: 2.6 angstroms
COMMENT R-value: 0.173
COMMENT Determination: X-ray diffraction
KEYWORDS hydrolase; serine proteinase
FEATURE
26-28, 35-38, 85-89,
69-73
#region beta sheet\
#region beta sheet\
136-141, 116-120,
179-183, 186-193,
204-209, 159-163,
144-152
#region beta sheet\
#region helix (right hand alpha) (single alpha turn, rest
irreg.)\
#region helix (right hand 3-10) (contiguous with h3)\
#region helix (right hand alpha) (contiguous with h2)\
#disulfide_bonds\
#active_site Asp, His, Ser #label CAT\
#site #label S13
#length 223 #molecular-weight 23071 #checksum 5569
SUMMARY
Query Match 39.5%; Score 768; DB 5; Length 223;
Best Local Similarity 46.5%; Pred. No. 3.60e-154;
Matches 106; Conservative 48; Mismatches 68; Indels 6; Gaps 6;
Db 1 IVGYTCGANTVPQVLSNGYHF-CGGSLINSQWVVAHCKYKGIQVRLGDNINVVE 59
Qy 33 VLGGHCQPHSQFQWQAALFQGLLQGLVGNWVLTAAHCKPKYTVRLGDHSLQNKD 92
Db 60 GNEQFISAKSIIVHPSYNSNTL-N-N-DIMLIKLSAASLNSRVASISLPTSCASAGTQC 116
Qy 93 GPEQIIPVQSIHPHCYNSSDVEDHNDMLQLRDQASLGSVKVPISLADHCTQPGKC 152
Db 117 LISGWGNTKSSGTSPDVLKCLKAPILSDSSCKSAVPQGITSNMFCAGLEGGDSQCGD 176
Qy 153 TVSGWGTVPSPRENFDTLNCAEVKIFPKQKCEDAYPGQITDGMVCAGS-SKGADTCQGD 211
Db 177 SGPVTVSGKLGIVSWGSG-CAAKNKPQVTKVCNYSWIKOTIASN 223
Qy 212 SGGPLVCDGALQGITSWGSDPCGRSKPGVYTNICRYLDWIKKIIGSK 259
RESULT 11
ENTRY B25528 #type complete
TITLE trypsin (EC 3.4.21.4) precursor - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change
20-Mar-1998
ACCESSIONS B25528
REFERENCE A93646
#authors
Stevenson, B.J.; Hagenbuechle, O.; Wellauer, P.K.
#journal
Nucleic Acids Res. (1986) 14:8307-8330
#title
Sequence organisation and transcriptional regulation of the
mouse elastase II and trypsin genes.

```

```

#cross-references MUID:87066713
#accession B25528
#molecule_type mRNA
#residues 1-246 #label STE
#cross-references GB:X04574; NID:g54918; PID:g54919
CLASSIFICATION #superfamily trypsin; trypsin homology
KEYWORDS calcium binding; hydrolase; protein digestion; serine
proteinase
FEATURE
1-23 #domain signal sequence #status predicted #label SIG\
24-246 #product trypsin #status predicted #label MAT\
24-239 #domain trypsin homology #label TRY\
30-160,48-64, #disulfide_bonds #status predicted\
132-233,139-206, #active_site His, Asp, Ser #status predicted\
171-185 #binding_site calcium (Glu, Asn, Val, Glu) #status
63,107,200 predicted
75,77,80,85
SUMMARY #length 246 #molecular-weight 26203 #checksum 5833
Query Match 39.5%; Score 768; DB 2; Length 246;
Best Local Similarity 42.7%; Pred. No. 3,60e-154;
Matches 106; Conservative 56; Mismatches 79; Indels 7; Gaps 7;
Db 5 LILALVGAA-VAFPVDDDKIVGGYTCRESSVPYQVSLNAGYHF-CGSLINDQWVYSA 62
QY 13 MFLLLGGAWAGHSRAQEDKVLGHECQPHSQPQAAALFQOQLCGGLVGGNWLTA 72
Db 63 HCYKRTQVRLGHNINVLGNEQFVDSAKIIRHPNNSWTL-D-N-DIMLIKASPVTL 119
QY 73 HCKRPKYTVRLGHSLOKNGDPEQIPVQSIHPHCYNSSDVEDHNDMLLQLRQASL 132
Db 120 NARVAVPLPSSCAPAGTQCLISGNGNTLSNGVNPDLQCVDPVLPQADCEASYPGDI 179
QY 133 GSKVKPISLADHCTQPGQKCTVSGWGTVPSPRENFDTLNCAEVKIFPQKKCEDAIPGQI 192
Db 180 TNNMICVGFLEGGKDCQGDGGGPPVNCGLQIVSMGYG-CAQDPAGVYTKVNTVDW 238
QY 193 TDGMVACAGSSKGA-DTCQGDGGPLVCDGALQGITSGSDPCGRSDRPGVYTNICRYLDW 251
Db 239 IQNTIADN 246
QY 252 IKKIIGSK 259

RESULT 12 TRRT1 #type complete
ENTRY trypsin (EC 3.4.21.4) I precursor - rat
TITLE trypsinogen I
FORMAL_NAME Rattus norvegicus #common_name Norway rat
REFERENCE 17-Dec-1982 #sequence_revision 17-Dec-1982 #text_change
20-Mar-1998
ACCESSION B22657 A00948
REFERENCE R.J.; Rutter, W.J.
#authors Craik, C.S.; Choo, Q.L.; Swift, G.H.; Quinto, C.; MacDonald,
J. Biol. Chem. (1984) 259:14255-14264
#journal Structure of two related rat pancreatic trypsin genes.
#cross-references MUID:85054880
#accession B22657
#molecule_type DNA
#residues 1-246 #label CRA
#cross-references GB:J00778; NID:g206507; PID:g206508
#note the authors translated the codon ATC for residue 6 as
Leu and GAC for residue 170 as Asn
REFERENCE A00948
#authors MacDonald, R.J.; Stary, S.J.; Swift, G.H.
#journal J. Biol. Chem. (1982) 257:9724-9732
#title Two similar but nonallelic rat pancreatic trypsinogens.
#cross-references MUID:82265624
#accession A00948
#molecule_type mRNA

```

```

#residues 1-246 #label MAC
GENETICS
14/1: 67/2: 152/1: 197/3
CLASSIFICATION #superfamily trypsin; trypsin homology
KEYWORDS hydrolase; pancreas; protein digestion; serine proteinase;
zymogen
FEATURE
1-15 #domain signal sequence #status predicted #label SIG\
16-23 #domain activation peptide #status predicted #label AP\
24-246 #product trypsin I #status predicted #label ENZ\
24-239 #domain trypsin homology #label TRY\
30-160,48-64, #disulfide_bonds #status predicted\
132-233,139-206, #active_site His, Asp, Ser #status predicted\
171-185 #binding_site calcium (Glu, Asn, Val, Glu) #status
63,107,200 predicted
75,77,80,85
SUMMARY #length 246 #molecular-weight 25959 #checksum 6732
Query Match 39.5%; Score 767; DB 1; Length 246;
Best Local Similarity 42.3%; Pred. No. 6,33e-154;
Matches 105; Conservative 56; Mismatches 80; Indels 7; Gaps 7;
Db 5 LILALVGAA-VAPPLEDDDKIVGGYTCPEHSVPYQVSLNAGYHF-CGSLINDQWVYSA 62
QY 13 MFLLLGGAWAGHSRAQEDKVLGHECQPHSQPQAAALFQOQLCGGLVGGNWLTA 72
Db 63 HCYKRTQVRLGHNINVLGDEQFVNAKIIKHPNYSWTL-N-N-DIMLIKASPVKL 119
QY 73 HCKRPKYTVRLGHSLOKNGDPEQIPVQSIHPHCYNSSDVEDHNDMLLQLRQASL 132
Db 120 NARVAVPLPSSCAPAGTQCLISGNGNTLSNGVNPDLQCVDPVLPQADCEAAVPGEI 179
QY 133 GSKVKPISLADHCTQPGQKCTVSGWGTVPSPRENFDTLNCAEVKIFPQKKCEDAIPGQI 192
Db 180 TNNMICVGFLEGGKDCQGDGGGPPVNCGLQIVSMGYG-CALPONPGVYTKVNFVWG 238
QY 193 TDGMVACAGSSKGA-DTCQGDGGPLVCDGALQGITSGWSDPCGRSDRPGVYTNICRYLDW 251
Db 239 IQNTIADN 246
QY 252 IKKIIGSK 259

RESULT 13 TRPS #type complete
ENTRY Trypsin complexed with inhibitor a90720a - bovine
TITLE #formal_name Bos primigenius taurus #common_name cattle
#note pancreas
REFERENCE A67200
#authors Lee, A.Y.; Clardy, J.
#submission submitted to the Brookhaven Protein Data Bank, September 1994
#cross-references PDB:1TPS
REFERENCE TN023754
#authors Lee, A.Y.; Smith, T.A.; Bonjouklian, R.; Clardy, J.
#journal Chem. Biol. (1994) 1:113
#title Atomic structure of the trypsin-a90720a complex: a unified
approach to structure and function.
COMMENT Resolution: 1.9 angstroms
COMMENT Determination: X-ray diffraction
KEYWORDS Hydrolase(serine protease)
FEATURE
144-156 #region helix (right hand alpha) (sngl alpha turn, rest
irreg.)\
208-213 #region helix (right hand 3-10) (contiguous with h3)\
214-223 #region helix (right hand alpha) (contiguous with h2)\
7-137 #disulfide_bonds\
25-41 #disulfide_bonds\
109-210 #disulfide_bonds\
116-183 #disulfide_bonds\
148-162 #disulfide_bonds\
173-197 #disulfide_bonds\
SUMMARY #length 223 #molecular-weight 23305 #checksum 7159

```

[illegible]

```

RESULT 15
ENTRY 1JRTA #type complete
TITLE trypsin (EC 3.4.21.4), chain A - bovine
PDB TITLE hemiacetal complex between leupeptin and trypsin
ORGANISM #formal_name Bos taurus; #common_name bovine
REFERENCE A65928
#authors Kurinov, I.V.; Harrison, R.W.
#submission submitted to the Brookhaven Protein Data Bank, February 1996
#cross-references PDB:1JRT
REFERENCE TN027917
#authors Kurinov, I.V.; Harrison, R.W.
#journal Nat. Struct. Biol. (1994) 1:735
#title Prediction of new serine proteinase inhibitors.
COMMENT Resolution: 1.8 angstroms
COMMENT Determination: X-ray diffraction
COMMENT R-value: no refinement
KEYWORDS digestion; hydrolase; pancreas; serine protease; zymogen
FEATURE
39-41 #region helix (right hand 3-10)\
145-151 #region helix (right hand alpha)\
209-211 #region helix (right hand 3-10)\
63-66 47-50,15-19,
23-31,34-37,86-90,
67-72 #region beta sheet\
136-141,115-120,
180-183,186-193,
204-208,160-163
7-137 #region beta sheet\
25-41 #disulfide_bonds\
109-210 #disulfide_bonds\
116-183 #disulfide_bonds\
148-162 #disulfide_bonds\
173-197 #disulfide_bonds\

SUMMARY
#length 223 #molecular-weight 23305 #checksum 7159

Query Match 39.4%; Score 766; DB 5; Length 223;
Best Local Similarity 46.5%; Pred. No. 1,11e-153;
Matches 106; Conservative 48; Mismatches 68; Indels 6; Gaps 6;

Db 1 IVGGYTCGANTVPYQVSLNSGYHF-CGGSLLNSQWVWSAAHCYKSGIOVRIGEDININVVE 59
QY 33 VLGGHCPCPSQPWQAALPQGOQLLCGGVLYGVGNWLVTAACHKPKYTVLRGDHSLQNKD 92
Db 60 GNEQFISAKSKIYHPHSYNNTL-N-N-DIMLKLSAASLSNRVASISLPTSCASAGTQC 116
QY 93 GPQEIPVTVQSPHPHCYNSSDVEDHNDMLQLQRQASLGSKVAPISLADHCTQPGQKC 152
Db 117 LISGWNKTSKSTSYPDVLKCLKAPILSDSSCKSAYPGQITSNMFCAVLEGGKSCQGD 176
QY 153 TVSGWGTVTSPRENFPDNLNCAEVKIFPQKKCEDAYPGQITDGMVCAGSSKA-DTCQGD 211
Db 177 SGPVVCSSKLOGIVSWGSG-CAQKNKPGVYTKVNCNVSWIKOTIASN 223
QY 212 SGGLPYCDGALOGITSWGSDPGGRSDPKGVYTNICRYLDWIKKIIGSK 259

```

Search completed: Wed Jul 22 15:28:54 1998  
Job time : 73 secs.

**THIS PAGE BLANK (USPTO)**



\*\*\*\*\*

MPARLH (TM)

Release 3.1A John F. Collins, BioComputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

on: Wed Jul 22 15:29:54 1998; MasPar time 15.26 Seconds  
Tabular output not generated. 717.292 Million cell updates/sec

Title: >US-08-915-659A-7  
Perfect Score: 1943  
Sequence: 1 MGRPRPRAAKTWMLLLGG.....VYTNICRYLDWIKKIGSKG 260

Scoring table: PAM 150  
Gap 11

Searched: 140542 seqs, 42109429 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: spiremb15  
1:sp\_fungi 2:sp\_human 3:sp\_invertebrate 4:sp\_mammal  
5:sp\_mmc 6:sp\_organelle 7:sp\_phage 8:sp\_plant  
9:sp\_bacteria 10:sp\_rodent 11:sp\_virus 12:sp\_vertebrate  
13:sp\_unclassified

Statistics: Mean 42.195; Variance 59.229; scale 0.712

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1488	76.6	260	10	SERINE PROTEASE INHIBI	0.00e+00
2	843	43.4	244	2	PROTEASE M PRECURSOR.	1.15e-206
3	776	39.9	273	2	NS1-NORMAL EPITHELIAL	5.59e-187
4	757	39.0	242	12	TRYPSIN PRECURSOR (EC	2.06e-181
5	754	38.8	237	12	TRYPSIN (FRAGMENT)	1.55e-180
6	751	38.7	247	12	TRYPSIN A2 PRECURS	1.17e-179
7	751	38.7	261	10	TISSUE KALLIKREIN.	1.17e-179
8	750	38.6	247	12	TRYPSIN A3 PRECURS	2.30e-179
9	744	38.3	245	12	TRYPSIN B2 PRECURS	1.31e-177
10	738	38.0	244	12	TRYPSIN B1 PRECURS	7.44e-176
11	710	36.5	259	2	TRYPSIN IV B-FORM.	1.12e-167
12	703	36.2	258	4	KALLIKREIN PRECURSOR.	1.23e-165
13	703	36.2	261	4	KALLIKREIN PRECURSOR	1.23e-165
14	694	35.7	239	10	KALLIKREIN (FRAGMENT).	5.14e-163
15	663	34.1	249	12	PREPROTRYPsin PRECURSO	5.33e-154
16	659	33.9	235	10	KALLIKREIN (FRAGMENT)	1.73e-153
17	644	33.1	219	12	TRYPSIN I (FRAGMENT)	7.73e-148
18	613	31.5	258	12	VENOM PLASMINOGEN ACTI	1.62e-139
19	593	30.5	262	12	CALOBIN.	9.53e-134
20	590	30.4	260	12	SERINE PROTEASE PRECUR	6.97e-133

21	581	29.9	258	12	013059	SERINE PROTEASE PRECUR	2.72e-130
22	579	29.8	260	12	013056	SERINE PROTEASE PRECUR	1.02e-129
23	576	29.6	250	10	003955	GLANDULAR KALLIKREIN (	7.46e-129
24	570	29.3	258	12	042207	CAPILLARY PERMEABILITY	3.96e-127
25	568	29.2	258	12	013063	SERINE PROTEASE PRECUR	1.49e-126
26	565	29.1	257	12	013057	MUCOFIRASE 3.	1.08e-125
27	560	28.8	260	12	013057	SERINE PROTEASE PRECUR	2.94e-124
28	557	28.7	257	12	013062	SERINE PROTEASE PRECUR	2.13e-123
29	552	28.4	257	12	013507	MUCOFIRASE 1.	5.79e-122
30	545	28.0	257	12	091508	MUCOFIRASE 2.	5.86e-120
31	545	28.0	257	12	091500	PREPROTRIMUBIN PRECURS	5.86e-120
32	538	27.7	257	12	091510	MUCOFIRASE 4.	5.90e-118
33	537	27.6	257	12	091511	MUCOFIRASE 5.	1.14e-117
34	520	26.8	257	12	013058	SERINE PROTEASE PRECUR	8.16e-113
35	521	26.8	263	10	035205	GRANZYME K.	4.23e-113
36	518	26.7	257	12	013069	KN-BJ2 PRECURSOR.	3.03e-112
37	509	26.2	258	10	061280	PRECURSOR ADIPSIN.	1.11e-109
38	504	25.9	248	10	061366	NATURAL KILLER CELL PR	2.94e-108
39	502	25.8	260	12	013061	SERINE PROTEASE PRECUR	1.09e-107
40	480	24.7	248	10	063224	GRANZYME-LIKE PROTEIN	1.91e-101
41	476	24.5	264	10	008643	LYMPHOCTYE MET-ASE 1 P	2.58e-100
42	476	24.5	271	3	018487	CHYMOTRYPSIN BI (FRAGM	2.58e-100
43	473	24.3	135	10	062284	NERVE GROWTH FACTOR, G	1.82e-99
44	471	24.2	225	2	Q15358	SERINE PROTEASE HOMOLO	6.71e-99
45	459	23.6	560	2	Q14520	HGF ACTIVATOR LIKE PRO	1.64e-95

ALIGNMENTS

RESULT	ID	PRELIMINARY	PRT	260 AA.
1	Q61955			
AC	Q61955			
DT	01-NOV-1996 (TREMREL. 01, CREATED)			
DT	01-NOV-1996 (TREMREL. 01, LAST SEQUENCE UPDATE)			
DT	01-JAN-1998 (TREMREL. 05, LAST ANNOTATION UPDATE)			
DE	SERINE PROTEASE INHIBITOR 5 (NEUROPSIN).			
GN	SP15			
OS	MUS MUSCULUS (MOUSE).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
OC	EUTHERIA; RODENTIA.			
RN	[1]			
RC	SEQUENCE FROM N.A.			
RC	STRAIN-BALB/C; TISSUE-HIPPOCAMPUS;			
RX	MEDLINE; 95348817.			
RA	ZU-LIN C., YOSHIDA S., KATO K., MOMOTA Y., SUZUKI J., TANAKA T.,			
RA	ITO J., NISHINO H., AIMOTO S., KIYAMA H.;			
RL	J. NEUROSCI. 15:5088-5097(1995).			
DR	EMBL; D30785; G1020091; -.			
DR	MGI; 104519; SP15.			
KW	SERINE PROTEASE.			
SQ	SEQUENCE 260 AA; 28523 MW; 232F78DB CRC32;			

Query Match	76.6%	Score 1488;	DB 10;	Length 260;
Best Local Similarity	72.2%	Pred. No. 0.00e+00;		
Matches	187;	Conservative 38;	Mismatches 34;	Indels 0; Gaps 0;
Db	1	MGRPPPCAQTQPIILLIFMGAWAGLTRAQSKILGREGICPHSQPQAAALFOGERLICGG	60	
Qy	1	MGRPRPRAAKTWMLLLGGAWAGHRAQEDKVLGHECQHPQAAALFOGQQLICGG	60	
Db	61	VLVGRWVLTAAHCKKQKYSVRLGHSLSQSRQPEQIQVAQSIQHPCYNNNPEDHSHD	120	
Qy	61	VLVGNWVLTAAHCKKPKYTVRLGHSLSQSRQPEQIPVQSIQHPCYNNNPEDHSHD	120	
Db	121	IMLRLONSANLGDKVKPVQLANLCPKVGOKCIIISGWCITVTSPOENFNTLCAEVIYS	180	
Qy	121	LMELQLRDQASLGSKVKPISLADHCTQPGQKCTVSGWGTVTSPTLCAEVIKIFP	180	
Db	181	QNKCRAPYPGKITEGMVCAAGSNGADTCOGDSGGPLVCDGLQGITSWGSDPCGKPEKG	240	
Qy	181	QKCEDAPYQGITDGMVCAAGSNGADTCOGDSGGPLVCDGLQGITSWGSDPCGKPEKG	240	
Db	241	VYTKICRYTTWIKTKMDNR	259	



Db 232 NDWL 235  
Qy 249 LDWI 252

## RESULT 5

ID Q91515 PRELIMINARY; PRT; 237 AA.  
AC Q91515;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DE TRYPsinogen (FRAGMENT).  
OS FUGU RUBRIPES (JAPANESE PUFFERFISH).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;  
OC OSTEICHTHYES; ACTINOPTERYGII; TETRAODONTIFORMES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA WANG K., GAN L., LEE I., ROACH J., HOOD L.;  
SUBMITTED (APR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
EMBL; U25747; G971196; -.  
FT NON\_TER 1  
SQ SEQUENCE 237 AA; 25726 MW; 6A296989 CRC32;

Query Match 38.8%; Score 754; DB 12; Length 237;  
Best Local Similarity 42.4%; Pred. No. 1.55e-180;  
Matches 103; Conservative 64; Mismatches 68; Indels 8; Gaps 8;

Db 1 LIAAAYAA-PIDEKIVGGYECKNSVAYQVSLNSGYHF-CGSLYNNVNSAAHCYK 58  
Qy 17 LGGAWAGHSRAQEDKVLGGHECQPHSQPWAALFQGGQLLCGGVGLVGNWLVTAACHCK 76  
Db 59 SRVVYRLGEHNRANEGPEQFISSSRVIRHPNYSYNI-D-N-DIMLIKSKPATLNOYV 115  
Qy 77 PKYTVRLGDHSLQNKDGPQEIIPVQSPHPHCYNSSDVEDHNDMLLQLRDQASLSKV 136  
Db 116 OPVALPSSCAAAGTKVYSGNGTMSSTAD-RNKLCNLNIPILSDRCENSYPGMITDAM 174  
Qy 137 KPISLADHCTQPGKCTVSGNGTVPSPRENFPTLNCAEVKIFFQKKCEDAYPGQITDGM 196  
Db 175 FCAGVLEGKSCQDGGPVPVCCNELQGVVSWGYG-CAERDHPGVYAKVCLFNDWLEST 233  
Qy 197 VCAGSSKGA-DTCQDGGPLVCDGALQGITSWGSDPCGRSDKPGVYTNICRYLDWIKKI 255

Db 234 MAS 236

Qy 256 IGS 258

## RESULT 6

ID O42158 PRELIMINARY; PRT; 247 AA.  
AC O42158;  
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
DE TRYPsinogen A2 PRECURSOR.  
GN TRYP2.  
OS PETROMYZON MARINUS (SEA LAMPREY).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES;  
OC AGNATHA (CYCLOSTOMATA).  
RN [1]  
RP SEQUENCE FROM N.A.  
RA ROACH J.C.;  
SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
EMBL; AF011898; G2367495; -.  
DR EMBL; AF011898; G2367495; -.  
KW SIGNAL.  
FT SIGNAL 1 15 POTENTIAL.  
FT CHAIN 16 247 TRYPsin A2.  
SQ SEQUENCE 247 AA; 26309 MW; D049AA4C CRC32;

Query Match 38.7%; Score 751; DB 12; Length 247;  
Best Local Similarity 42.0%; Pred. No. 1.17e-179;  
Matches 105; Conservative 59; Mismatches 77; Indels 9; Gaps 9;

Db 4 LIALLVGVAAPYMYEDHIVGSECAAHSPQWQVSLNIGYHF-CGSLINSQWVYSA 62  
Qy 13 MFLILLGGAWAGHSRAQEDKVLGGHECQPHSQPWAALFQGGQLLCGGVGLVGNWLVTA 72  
Db 63 HCYQTASRISVRIGEHNFVNEGTEQIOAKAIOHPYNSWTI-D-N-DIMLIKLSPPA 119  
Qy 73 HC-KKP-KYTVRLGDHSLQNKDGPQEIIPVQSPHPHCYNSSDVEDHNDMLLQLRDQA 130  
Db 120 TLNOYAAIAIALPSSCVNTGVMCTISGWEQTTSIGS-PDVLMCVQAPVLSDTSCRSNYPG 178  
Qy 131 SLGSKVAPISLADHCTOPGOKCTVSGNGTVPSPRENFPTLNCAEVKIFFQKKCEDAYPG 190  
Db 179 DITNNMTCILGLEKDKSCQDGGPVPVCCNELQGVVSWGYG-CALPNYPGVYTKVCNIN 237  
Qy 191 QITDGMVCASSKGA-DTCQDGGPLVCDGALQGITSWGSDPCGRSDKPGVYTNICRYL 249  
Db 238 AWTAAQTAAAN 247  
Qy 250 DWIKKIIGSK 259

## RESULT 7

ID Q61855 PRELIMINARY; PRT; 261 AA.  
AC Q61855;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DE TISSUE KALLIKREIN.  
GN MKR-6.  
OS MUS MUSCULUS (MOUSE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA TADA M., PETERS J., TAKAHASHI S., INOUE H., MIYAKE Y.;  
SUBMITTED (AUG-1994) TO EMBL/GENBANK/DBJ DATA BANKS.  
EMBL; D10464; G520591; -.  
SQ SEQUENCE 261 AA; 28775 MW; 509D1D00 CRC32;

Query Match 38.7%; Score 751; DB 10; Length 261;

Best Local Similarity 42.4%; Pred. No. 1.17e-179;

Matches 108; Conservative 57; Mismatches 78; Indels 12; Gaps 9;

Db 6 LFLALSLLGGI-DAAPPVQ-SRIVGNCENKSNOPQWAVYRFTKYQCGGILLNANWVITA 63  
Qy 13 MFLLL-LGGAWAGHSRAQEDKVLGGHECQPHSQPWAALFQGGQLLCGGVGLVGNWLVTA 71  
Db 64 AHCNDKYQVWLKKNFLEDEPSAQHRLVSKAIPHPDNNMSSLNEHTPQPEDDYSDNML 123  
Qy 72 AHCKPKYTVRLGDHSLQNKDGPQEIIPVQSPHPHCYN-S--SD-V--EDH-NHDLML 123  
Db 124 LRLKPPADITDVVKPIDLPTEEPKLGSTCLASGWSITPVKYEYDDELQCVNKLKLPED 183  
Qy 124 LQLRDQASLGSKVKPISLADHCTQPGOKCTVSGNGTVPSPRENFPTLNCAEVKIFFQKK 183  
Db 184 CAAHIEKVTDDMLCAGDMGDKTACAGDSGGLPCDGLQGITSWGSPGKPNVPGIY 243  
Qy 184 CEDAYPGQITDGMVCASSKGA-DTCQDGGPLVCDGALQGITSWGSDPCGRSDKPGVY 242  
Db 244 TRVLNFTWIRETMA 258  
Qy 243 TNICRYLDWIKKIIG 257

## RESULT 8

ID O42608 PRELIMINARY; PRT; 247 AA.  
AC O42608;  
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
DE TRYPsinogen A3 PRECURSOR.  
GN TRYP3.

OS PETROMYZON MARINUS (SEA LAMPREY).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES;  
 OC AGNATHA (CYCLOSTOMATA).  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA ROACH J.C.;  
 RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=ANTERIOR INTESTINE;  
 RA ROACH J.C.;  
 RL SUBMITTED (AUG-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 DR EMBL; AF011899; G2367497;  
 DR EMBL; AF011352; G2293478;  
 KW SIGNAL.  
 FT SIGNAL  
 FT CHAIN  
 FT CHAIN  
 SQ SEQUENCE 247 AA; 26295 MW; 5071AC2E CRC32;

Query Match 38.6%; Score 750; DB 12; Length 247;  
 Best Local Similarity 42.0%; Pred. No. 2.30e-179;  
 Matches 105; Conservative 59; Mismatches 77; Indels 9; Gaps 9;  
 4 LIALLVGAAPYEDHIVGSECAHSPQWVSLNIGYHF-CGGSLINSQWVSA 62  
 QY 13 MFLLLGGAWAGHRAQEDKVLGGHECPHSPQWQAALFQOQLCGVLYGGNVLTA 72  
 Db 63 HCQTASRISVRIGEHNFVTEGTQRIQASKAIRHPQYNSATI-D-N-DIMLIKSSPA 119  
 QY 73 HC-KKP-KYTVRLGDHSLQNKDGPQEIIPVQSIHPHCYNSSDVEDHNDMLQLRDQA 130  
 Db 120 TLNQYQAQVLPSSCVGTGVMCTISGGETQTSVGS-PDVLVCVQAPVLSDTSCRSNYPG 178  
 QY 131 SLGSKVKPISLADHCTOPGQKCTVSGWGTVTSPRENFDTLNCAEVKIFPQKKCEDAYPG 190  
 Db 179 DITNNMICLGLYEGKDCSGDGGPVVNCGLQIGIVSWGRG-CALPNYPGVYTKVCN 237  
 QY 191 QITDGMVCAAGSSKGA-DTCQDGGPLVCDGALQGITSWGSDPCGRSKDPGVYTNICRYL 249  
 Db 238 AWIAQTIAAN 247  
 QY 250 DWIKKIIGSK 259

RESULT 9  
 ID 042160 PRELIMINARY; PRT; 245 AA.  
 AC 042160;  
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
 DE TRYPB2.  
 GN TRYPB2.  
 OS PETROMYZON MARINUS (SEA LAMPREY).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES;  
 OC AGNATHA (CYCLOSTOMATA).  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA ROACH J.C.;  
 RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 DR EMBL; AF011901; G2367501;  
 KW SIGNAL.  
 FT NON\_TER 1 1  
 FT SIGNAL <1 13 POTENTIAL.  
 FT CHAIN 14 245 TRYPB2.  
 SQ SEQUENCE 245 AA; 26001 MW; 071872F0 CRC32;

Query Match 38.3%; Score 744; DB 12; Length 245;  
 Best Local Similarity 42.0%; Pred. No. 1.31e-177;  
 Matches 105; Conservative 61; Mismatches 74; Indels 10; Gaps 10;  
 3 IFALLVGTGA-AAAPYMWEDHIVGSECAHSPQWVSLNIGYHF-CGGSLISSEWVSA 60  
 QY 13 MFLLLGGAWAGHRAQEDKVLGGHECPHSPQWQAALFQOQLCGVLYGGNVLTA 72

Db 61 HCQTASRISVRIGEHNFVTEGTQRIQASKAIRHPQYNSATI-D-N-DIMLIKSSPA 117  
 QY 73 HC-KKP-KYTVRLGDHSLQNKDGPQEIIPVQSIHPHCYNSSDVEDHNDMLQLRDQA 130  
 Db 118 TLNQYQAQVLPSSCVGTGVMCTISGGETQTSVGS-PDVLVCVQAPVLSDTSCRSNYPG 176  
 QY 131 SLGSKVKPISLADHCTOPGQKCTVSGWGTVTSPRENFDTLNCAEVKIFPQKKCEDAYPG 190  
 Db 177 DITNNMICLGLYEGKDCSGDGGPVVNCGLQIGIVSWGRG-CALPNYPGVYTKVCN 235  
 QY 191 QITDGMVCAAGSSKGA-DTCQDGGPLVCDGALQGITSWGSDPCGRSKDPGVYTNICRYL 249  
 Db 236 SWIASTMAAN 245  
 QY 250 DWIKKIIGSK 259

RESULT 10  
 ID 042159 PRELIMINARY; PRT; 244 AA.  
 AC 042159;  
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
 DE TRYPB1.  
 GN TRYPB1.  
 OS PETROMYZON MARINUS (SEA LAMPREY).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES;  
 OC AGNATHA (CYCLOSTOMATA).  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA ROACH J.C.;  
 RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 DR EMBL; AF011900; G2367499;  
 KW SIGNAL.  
 FT NON\_TER 1 1  
 FT SIGNAL <1 12 POTENTIAL.  
 FT CHAIN 13 244 TRYPB1.  
 SQ SEQUENCE 244 AA; 25903 MW; 41F78768 CRC32;

Query Match 38.0%; Score 738; DB 12; Length 244;  
 Best Local Similarity 41.2%; Pred. No. 7.44e-176;  
 Matches 103; Conservative 63; Mismatches 74; Indels 10; Gaps 10;  
 2 IFALLVGTGA-AAAPYMWEDHIVGSECAHSPQWVSLNIGYHF-CGGSLISSEWVSA 59  
 QY 13 MFLLLGGAWAGHRAQEDKVLGGHECPHSPQWQAALFQOQLCGVLYGGNVLTA 72  
 Db 60 HCQTASRISVRIGEHNFVTEGTQRIQASKAIRHPQYNSATI-D-N-DIMLIKSSPA 116  
 QY 73 HC-KKP-KYTVRLGDHSLQNKDGPQEIIPVQSIHPHCYNSSDVEDHNDMLQLRDQA 130  
 Db 117 TLNQYQAQVLPSSCVGTGVMCTISGGETQTSVGS-PDVLVCVQAPVLSDTSCRSNYPG 175  
 QY 131 SLGSKVKPISLADHCTOPGQKCTVSGWGTVTSPRENFDTLNCAEVKIFPQKKCEDAYPG 190  
 Db 176 DITNNMICLGLYEGKDCSGDGGPVVNCGLQIGIVSWGRG-CALPNYPGVYTKVCN 234  
 QY 191 QITDGMVCAAGSSKGA-DTCQDGGPLVCDGALQGITSWGSDPCGRSKDPGVYTNICRYL 249  
 Db 235 SWIASTMAAN 244  
 QY 250 DWIKKIIGSK 259

RESULT 11  
 ID Q15655 PRELIMINARY; PRT; 259 AA.  
 AC Q15655;  
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
 DE TRYPB1.  
 GN TRYPB1.  
 OS HOMO SAPIENS (HUMAN).



Query Match	34.1%	Score 663;	DB 12;	Length 249;
Best Local Similarity	41.8%	Pred. No. 5,338-154;		
Matches 104;	Conservative 49;	Mismatches 86;	Indels 10;	Gaps 8;
6	LLLLIGAAAAPR-EDGRIIGGVECSPHSRPMASLNYGHF-CGGVLLNNOWLWLSVAHC	63		
	: : : : :       : : : : :       : : : : :       : : : : :			
15	LLLLGGAWAGHSAQEDKVLGGHECOPHSQPQWAALFQQQLLGGGVLVGGNWVLTAHC	74		
	: : : : :       : : : : :       : : : : :       : : : : :			
64	WNPYMSQVILGDHNRVPEGTETQLMKNTNTI IWHPSYDYQTL-DF--DIMLIKLYHPVEV	120		
	:       : :       : :       : :       : :       : :       : :			
75	KRPKYT--VRLGDHSLQNDGPEQEIPVQVQSIPHPGPNCSDDVEDHNDMLLQLLRQASL	132		
	: :       : :       : :       : :       : :       : :       :			
121	TEAVAPIPLPTSCPYGGLSCSVGGIACLGEAYMPTLLQCLNVPYIDVOQCENTYPLC	180		
	: :       : :       : :       : :       : :       : :       :			
133	GSKVRFISLADHCTQPGQCKTVSGWGVTSPRENF--PDTLNCIAEYKIFPKKCEADYAPQ	191		
	: :       : :       : :       : :       : :       : :       :			
181	IISTVMVACYMEGGKDACNGDSGLPVCDEYQGLVSWG--QGCABPNYPGVYVKLCFEHS	239		
	: :       : :       : :       : :       : :       : :       :			
192	ITDGMVACGSSKGA-DTCQDGGGLPVCALQGITISWGDSPCGRSKDKPGVNTICRYLD	250		
	: :       : :       : :       : :       : :       : :       :			